INTEGRATING AVHRR DERIVED NDVI WITH ECOLOGICAL MODELLING

Jian Chen
Department of Geography
Rutgers University
New Brunswick, NJ 08903

ABSTRACT Annual net primary productivity (NPP) was found well correlated with integrated normalized difference vegetation index (NDVI) at a scale of major vegetation biomoes of the Continent. The correlation was physically bridged by the causal relationship between leaf area index (LAI) and absorbed photosynthetically active radiation (APAR). However validations of these models were often based on very limited experiment sites. In this project, an improved method of calculating integrated NDVI was proposed. Comparison with the integrated NDVI using conventional method revealed that much higher correlations existed between NPP and integrated NDVI estimated by the proposed method than that by conventional method for the forest in the Northeastern USA. Correlation analysis also revealed that for all forest types mean integrated NDVI and NPP calculated at forest patch scale showed much higher correlation compared to the result based on pixel by pixel comparison. This showed NDVI and NPP had higher agreement at forest patch scale by alleviating the noise brought in the analysis. The analysis was conducted by using ARC/INFO™, GRASS and S-PLUS™ software packages.

I. INTRODUCTION

Annual net primary production (NPP) from literature (Whittaker and Likens, 1975) was found well correlated with integrated NDVI at a scale of major vegetation biomes of the Continent (Goward et al., 1985). Running and Nemani (1988) related seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of forests in different climates. The correlation was physically bridged by the causal relationship between leaf area index (LAI) and absorbed photosynthetically active radiation (APAR) (Asrar et al., 1991). Montieth (1977) provided an example of estimating NPP solely from APAR rather than by using a carbon budget approach. Prince (1991) put forward a conceptual model of regional production for use with coarse resolution AVHRR data based on a model of crop primary production. In the last decade, there was an increasing number of ecological models estimating NPP through both lumped environment parameters and processes based approaches (Raich et al., 1991, Aber et al., 1991, Aber and Federer, 1992, Running and Hunt, 1993). However validations of these models were often based on very limited experiment sites and at a limited scale. AVHRR derived NDVI could be a valuable resources for extrapolating these ecological models to a broad scale (Prince, 1991, Jarvis, 1993, Reynolds et al, 1993).

The objectives of this project are (i). provide a more reasonable method of estimating integrated NDVI; and (ii). explore the relationship between integrated NDVI and annual NPP output from PnET for the Northeastern US forest.

II. DATA AND METHODOLOGY

2.1 Data Characteristics

The AVHRR data set distributed by the U.S. Geological Survey's EROS data center is composed of twenty-one two-week maximum NDVI composites for the conterminous U.S. These composites were generated from nearly 500 NOAA-11 daily images. The 17 core composite periods represent a continuous period from March 1, 1991 to October 22, 1991, which was intended to capture the characteristics of growing season in the most areas. The first two and last two composites represent a two-week period each for January, February, November and December. Each 14-day composite includes 10 bands of information. They are AVHRR channels 1-5 reflectance, NDVI, satellite zenith, solar zenith, relative azimuth, and date of observation. The data layers are extracted from the daily observation scene on the basis of the maximum NDVI compositing process. Each band is of dimension 2889 lines and 4608 samples and is geometrically registered to the Lambert Azimuthal Equal Area map projection of 1 kilometer resolution. The NDVI and date layers were utilized in the current project.

Annual NPP image for the Northeastern USA was obtained by running PnET model based on 30-year average climate condition. PnET is a simple, lumped parameter, monthly time step model of carbon and water balance of forests. The two principles driving the model are that maximum photosynthetic rate is a function of foliar nitrogen concentration and that stomatal conductance is a function of realized photosynthetic rate (Aber and Federer, 1992). The NPP output was provided in latitude and longitude coordinates for the Northeastern USA at a resolution of 30 seconds.

Land use/cover image for the Northeastern USA was also in latitude and longitude coordinates of resolution 30 seconds based on both 1989 AVHRR image and USGS land use/cover classification (Lathrop, 1994, personal communication). There were 11 land use/cover classes identified in the classification. Since the PnET model is a forest ecosystem model and generates annual NPP for forest area, five forest types were engaged in the current analysis. They were northern hardwoods, spruce-fir, northern NE mixed forest, mid-Atlantic forest, southern NE mixed forest/cropland/pasture.

Nineteen climatological stations were selected in the northeastern USA from the Weekly Weather and Crop Bulletin (1991) to estimate the starting and ending date of the growing season at different locations. The Sunday of the week in which there is no single day with extreme minimum temperature below zero Celsius in the Spring was selected as the beginning of the growing season. The Sunday of the first week in the fall in which there is at least one day of below zero Celsius was selected as the end of the growing season. The starting and ending date of the growing season was interpolated over the whole region based on the observation of these nineteen stations.

2.2 Data Processing and Methodology

Annual NPP image and land use/cover image were projected to Lambert Azimuthal Equal Area map projection using arc/info GRID package PROJECT module. 21 pairs of conterminous USA NDVI and date images were subset using a short C program to generate subimage for the Northeastern area. Since the date layer provided a pointer to the daily AVHRR images, actual observation date for each pixel was recorded as Julian day of the year. All the subset images were converted to grids in arc/info in the Lambert Azimuthal Equal Area projection.

Integrated NDVI were estimated using two methods. Conventional method assumes that composite NDVI is a representative for the whole two-week period and the starting and ending date for calculation of integrated NDVI is uniform over the whole region. In this case, the integrated NDVI was estimated by summing all the core 17 images (March 1 to October 22, 1991), dividing by two and multiply by the number of days in the whole period (196 days), which is integrated NDVI days -- a concept suggested by Goward et al. (1988) to avoid confusion of integrated NDVI by using different period composite images. This integrated NDVI days concept provided the comparability if one uses two-week composite images and the other use monthly composite images.

The proposed method adopted the integrated NDVI days concept and used the date information provided in the data set and the starting and ending date of the growing season interpolated from climatological stations. For each pixel the NDVI was plotted against the actual date of the NDVI observation. This approach did not assume that the date of the NDVI was in the middle of the two-week composite period. The area bounded by the starting and ending date of growing season and under the NDVI curve was actually calculated (Fig. 1).

The relationship between integrated NDVIs (using both methods) and annual NPP were examined both for all forest region and for "single" forest type area. At both scales, the relationship were explored on a pixel by pixel basis and small clump-mean basis. In the later case, single forest clumps of size ranging from 50-500 square kilometer were first identified (see discussion for detail), the mean value of NDVIs and NPP were calculated for each clump using ZONALMEAN module in arc/info. A comparison of correlations between NPP and the two integrated NDVIs were carried out. This process was made possible by output data from arc/info GRID using SAMPLE module as an ASCII file and then imported into SPLUS statistical package by using MATRIX and SCAN functions.

III. RESULTS AND DISCUSSIONS

By visually examining the NDVI composite images for the different periods, the green and brown wave phonology was clear in this region. This was well documented on the selected NDVI variation profiles for the northern, middle and southern portion of this region (Fig. 1). The green wave started at the middle of April at the southern portion, propagated northward and ended at the middle or late of May. The brown wave started at the middle September and pushed southward to the southern portion at early or middle October.

Integrated NDVIs estimated by different methods appear to have similar pattern at first look. However, careful examination found that integrated NDVI estimated by proposed method showed higher degree of latitudinal variation, which is in agreement with NPP variation trend in general in this region.

Taking each pixel as a sample, plots of NPP vs the integrated NDVI with conventional method and NPP vs the integrated NDVI using proposed method were generated and were subjected to regression analysis. All plots showed some linear trend, except for the plots for spruce-fir and northern NE mixed forest. However the plots were scattered, which was well documented by least square fit R-squares in Table 1. It was obvious that the proposed integrated NDVI were better correlated with NPP for all forest types. For spruce-fir, there was no correlation between NPP and integrated NDVIs. In the case of northern NE mixed forest, the plot seems to be a composite of linear trend and non-linear spruce-fir formation, which was coincide with the fact that northern NE mixed forest composed of hardwood and spruce-fir.

Table 1. Correlation between integrated NDVIs and NPP on pixel by Pixel basis

land cover R-square	whole region	Hardwood	Spruce-fir	Northern Mixed	Mid-Atlantic Mixed	Southern Mixed
NPP vs conventional NDVI	0.13	0.25	0.01	0.16	0.25	0.09
NPP vs Proposed NDVI	0.35	0.53	0.08	0.38	0.45	0.33
Sample Numbers	7505*	34950	12183	52927	21324	28399

*For the whole region, sampling was performed at every 5 km since SPLUS can not deal with sample file of size larger that 5 megabytes.

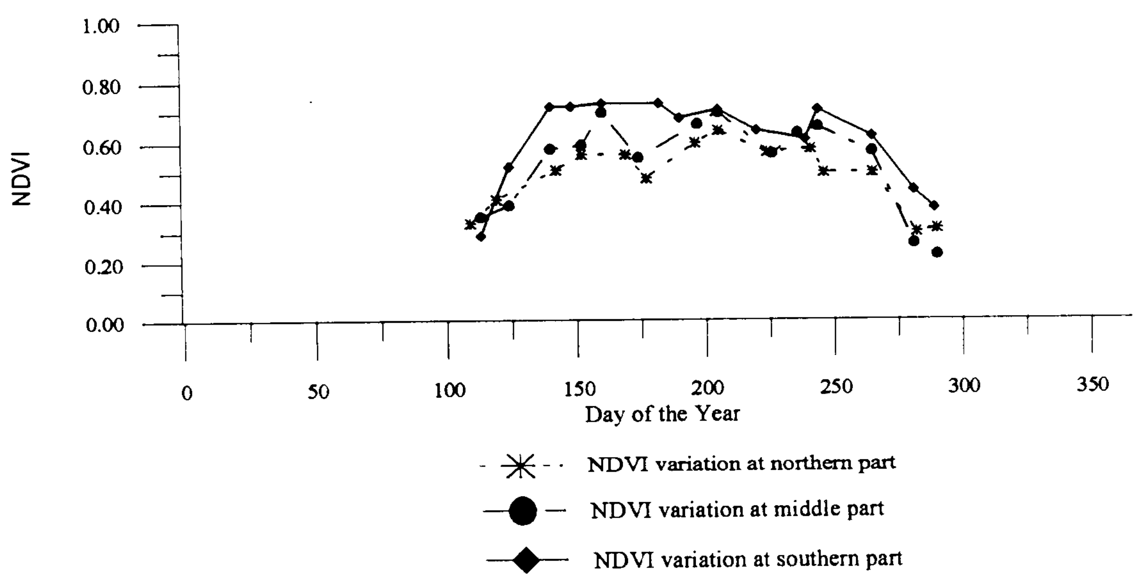
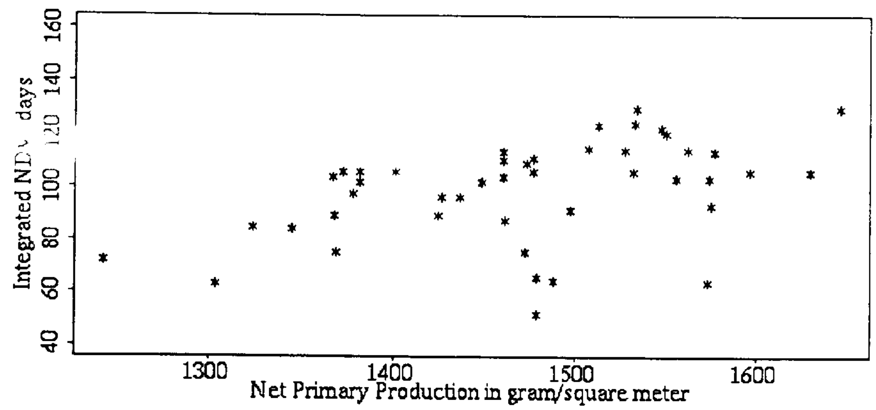


Figure 1. Typical NDVI variations at northern, middle and southern Part of the Northeastern US.



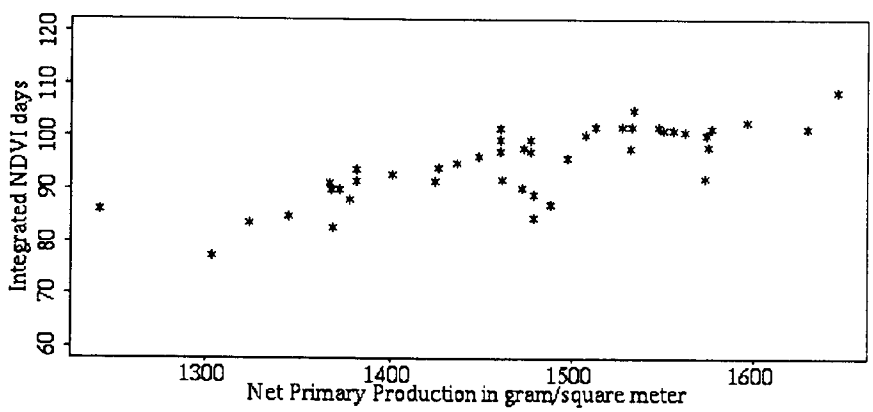


Figure 2. Relationship between NPP and integrated NDVI days for hardwood forest.
Top: integrated NDVI days were calculated by conventional method.
Bottom: integrated NDVI days were calculated by the proposed method.

All the forest region were clumped within single forest type using four-pixel contiguous criteria. Medium size clumps ranging from 50 to 500 pixels were selected for clump-mean NPP vs clump-mean integrated NDVIs analysis. The determination of the clump size was based on the fact that this size is big enough to sieve sporadic small clumped pixels and not too big to include pixels of large latitudinal range, so that the local information can be preserved (Fig 3). The selection also took the areal distribution of clumps into consideration.

Correlations on the clump-means basis (Table 2) also demonstrated that the proposed integrated NDVI were better correlated with NPP, which were supported by visual examination of scatter plots of NDVIs vs NPP. The plots of NPP vs proposed integrated NDVI tend to have steep slope and less scattered distribution compared to the plots of integrated NDVI with conventional method and NPP (for example Fig 2). R-square values clearly supported the statement.

Figure 3. Medium size forest clumps (50 - 500 pixels) identified in clump-mean NPP and integrated NDVI analysis.
Top left: Boundary of Northeastern US. Top middle: Hardwood clumps.
Top right: Spruce-fir forest clumps.
Bottom left: Northern NE mixed forest clumps.
Bottom middle: Mid_Atlantic mixed forest clumps.
Bottom right: Southern mixed forest clumps.

Table 2. Correlation between integrated NDVIs and NPP on the clump-mean basis

land cover R-square	whole region	Hardwood	Spruce-fir	Northern Mixed	Mid-Atlantic Mixed	Southern Mixed
NPP vs conven-tional NDVI	0.16	0.23	0.11	0.23	0.25	0.02
NPP vs Proposed NDVI	0.45	0.64	0.11	0.47	0.47	0.48
Sample Numbers	26023	5843	3302	7954	3360	5628

Of all the forest types, hardwood always has the highest correlation between integrated NDVI and NPP in all the cases, northern mixed forest and mid Atlantic forest have moderate correlation between NDVI and NPP. Southern mixed forest has lower correlation. Spruce-fir forest shows no correlation.

Prince (1991) claimed that Montieth's model (1972) based on crop yield should be applied to total assimilation in the following form:

$$Pg = Eg * sum(Nt * St) \qquad \text{..........} \quad (1)$$

where Pg is total assimilation. Eg is a measure of the dry-matter yield of APAR in total assimilation. Nt is the NDVI at time period t. St is the global photosynthetically active radiation incident during time interval t.

Montieth (1972), Jarvis and Leverenz (1983), Charles-Edwards et al. (1986) and Prince (1991) advocated the model using the terminology of Jarvis and Leverenz (1983):

$$Pn = f * Eg * Yg * Ym * d * sum(Nt * St) \qquad \text{..........} \quad (2)$$

where Pn is the net production, f measures the departure from maximum efficiency caused by physiological adjustment such as stomatal closure and uncoupling of the light and dark reaction of photosynthesis, Ym is the proportion of assimilate not used in maintenance respiration, Yg is a coefficient measuring the efficiency of conversion of assimilate into growth, and d is the proportion of growth not lose in death, decay and grazing (Prince, 1991).

Equation (2) well explained why the relationship between NDVI and cropland is generally better than that in forest area. Since these cropland usually have single type of vegetation and these coefficients mentioned could be approximately assumed constant. So linear relationship is expected between Pn and sum of Nt by assuming St also constant. This also explained the fact that correlation between NPP and integrated NDVIs for single forest type were generally higher than that for the whole forest region.

The correlation analysis conducted in this project seems further support the model presented in Equation (2). Hardwood has the highest correlation between NPP and integrated NDVI probably because of its simplicity of forest composition or due to the fact that the species have similar coefficient. Southern NE mixed forest has lower correlation because it includes wide range of cropland and pastures. Northern and mid-Atlantic mixed forest have moderate complexity and moderate correlation. Spruce-fir forest has little or no correlation probably due to three factors:

(i). Spruce-fir keep their leaves for up to 8 years (Ollingner, 1994, personal communication) so that Equation (2) is not appropriate for this forest type modeling.

(ii). There is a random generating process in PnET to estimate the percentages of spruce and fir in a particular pixel. So that the NPP estimated from a particular pixel or a particular local area may not be the actual condition for that area, while on the scale of the whole northeastern USA region the NPP is close to the regional average.

(iii). spruce and fir have different coefficients presented in Equation (2).

It was observed that the clump-mean integrated NDVI and NPP have higher correlation compared to that obtained in the pixel by pixel analysis. This may be due to the following facts.

(i). clump-mean is a closer representation of the average condition of the local area.

(ii). noise brought by sporadic pixels were eliminated.

(iii). modifiable area unit problem may come into play, but this is fair in this case because we are interested in much broader scale variation (the largest clump is about 23 by 23 kilometers)

(iv). using clump-means may alleviate the misregistration problem, if any, during the overlay of different images. There were 17 pairs NDVI and date images used in this analysis. Although it was claimed that these images were registered according to USGS DLG water body features and had RMSE (root mean square error) within a pixel, it was expected that misregistration existed in the process. Land use/cover classification used for PnET and in this project was based on USGS generated classification and 1989 AVHRR image. USGS classification was originally in UTM coordinate system. It was expected that errors associated with map projection conversion using nearest method in arc/info to convert NPP and land use/cover images from latitude and longitude to Lambert Equal Area projection.

IV. CONCLUSION AND FURTHER RESEARCH DIRECTIONS

Annual net primary production (NPP) was found fairly correlated with integrated NDVI for "single" forest types in the Northeastern USA. Hardwood forest showed the highest correlation compared to other forest types, northern mixed forest, mid-Atlantic mixed forest had moderate correlation, southern mixed forest showed lower correlation. There was little correlation between integrated NDVI and NPP for spruce-fir forest type. The proposed method of estimating integrated NDVI avoided errors in estimation due to irregular sampling space which was not uncommon in this data set. More significantly, by taking the starting and ending date of the growing season into consideration, the integrated NDVI was a closer representative of the actual phenological process in the ecosystem and the annual NPP. Correlation analysis revealed that much higher correlation between NPP and integrated NDVI estimated by the proposed method compared to that by conventional method for all forest types. Clump-means correlation analysis showed higher correlation between NPP and integrated NDVI by alleviating the noise brought in the analysis processes. This project further support the model proposed originally by Montieth (1977) and advocated by Jarvis and Laverenz (1982) and Prince (1991). Furthermore, this project demonstrated that estimation of the starting and ending date of a growing season should be considered as an important factor in calculating integrated NDVI.

The author also realized that there were certain aspects for improvement in relating integrated NDVI with NPP. By ignoring St in Equation (2) probably introduced considerable error in this region since the radiation variation over the latitudinal range could be important in this region. Further research may be conducted by incorporating radiation variation as a modifier in calculating integrated NDVI. In the current project, 19 weather stations were included in the process to estimate starting and ending date of the growing season for the whole region. Elevation influence most probably was not well represented. More station could be considered or the elevation could serve as another variable in the correlation analysis.

Bibliography

Aber, J. D. and Federer, C. A. 1992. A generalized, lumped-parameter model of photosynthesis, evaporation and net primary production in temperate and boreal forest ecosystems. Oecologia 92:463-474.

Goward, S. N., Tucker, C. J. and Dye, D. G. 1985. North American vegetation pattern observed with the NOAA-7 advanced very high resolution radiometer. Vegetatio 64:3-14.

Jarvis, P. G. and Laverenz, J. W. 1983. Productivity of temperate. deciduous and evergreen forest. In: Physiological Plant Ecology IV. Ecosystem Processes: Mineral Cycling, Productivity and Man's Influence. ed. O. L. Lange, P. S. Nobel, C. B. Osmond and H. Ziegler. New York:Academic Press.

Jarvis, P. G. 1993. Prospects for bottom-up models. In: Scaling Physiological Process: Leaf to Global. ed. J. R. Ehleringer and C. B. Field. New York: Academic Press.

Montieth, J. L. 1972. Solar radiation and productivity in tropical ecosystem. Journal of Applied Ecology. 9:747-766.

Montieth, J. L. 1977. Climate and the efficiency of crop production in Britain. Philosophical Transaction of the Royal Society 281:277-294

Prince, S. D. 1991. A model for regional primary production for use with coarse resolution satellite data. International Journal of Remote Sensing. 12:1313-1330.

Raich, J. R., Rastetter, E. B., Mellilo, J. M., Kicklighter, D. W., Steudler, P. A. and Peterson, B.J. Grace, A. L. Moore B. III and Vorosmarty, C. J. 1991.Potential net primary productivity in South America: application of a global model. Ecological Application 4:399-429.

Reynolds, J. F., Hilbert, D. W., Kemp, P. R. 1993. Scaling ecophysiology from the plant to the ecosystem: a conceptual framework. In: Scaling Physiological Process: Leaf to Global. ed. J.R.

Ehleringer and C. B. Field. New York: Academic Press. Running, S. W. and Nemani, R. R. 1988. Relating seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of forests in different climates. Remote Sensing of Environment 24:347-367.

Running, S. W. and Hunt, E. R. Jr. 1993. Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models. In: Scaling Physiological Process: Leaf to Global. ed. J. R. Ehleringer and C. B. Field. New York: Academic Press.

Whittaker, R. H. and Likens, G. E. 1975. Primary production: the biosphere and man. In: Primary Productivity of the biosphere. ed. H. Leith and R. H. Whittaker. Ecological Studies 14. New York: Spring Verlag.